Figure 1A Nucleotide sequence of inserted environmental DNA (029cel)
SEQ ID NO:1

```
ATCAACACGC TGGAAAGTAA TTTCAAGGGT AAGGCCATCG GTTGCCGCCG      50
GGGTAGAAAT GTGCGGTTGG ATTTCGTTGA GCGGCGTCGC CGGCGTTCCA     100
CCGAGGGCAT AGCGCAGCAG GTTGGCGATG CCACCGGTGA GGCCTTCGGG     150
GCCGCCTACG ATGTTGTGCT CAGCCGCCCA TGCGATGTAG CCGTCCGGCT     200
CGGGTTCGCT CGCGGGGGTG AAGAAGACAA TGTCGTCGAG ATAAAGGTTG     250
CCGCTTCCGC TCTCAACGCC GCCGAGGTTG AATTGGATTT CGCAAATTCT     300
CGTTAGGTCC AGCACGGAAT CGCCGACGAG GTCGGCTATG GGAATCTGAA     350
TGCGCCCATA GGGTTGGGTA CGCGGAAGGG ACACGTAGGG ACCCACTTTG     400
TCATTGGGCG AGACGAGCCG GACAAAGATT TGGTGCGCCG CCTGCGAGGG     450
GCCTTGGAGG GCGAGAGAAA GGTACGTGAG GGCGCTGATG TCGTGCGTGG     500
GACCGTCTCC CCAGTTGTCG AGATTGAGCC CAAATCCGGC CCACCATCCG     550
GCGATAGTGT AGCTCCAATG GTAGTGACGC TCACCCTCGA AGCCGCCGCT     600
GGAGAGTTCC TGCAAGCCGT CGCCCCAAAT GCCCGTGATG AGCGTTGCCT     650
CGTCACGGTA GATCACAAGT TCGGCGGCGG GTGCCGGGGG AAGATCGCCT     700
TGAGTGATCA CGAGAGTGGC GGTGGCGCTG CCTTCGTGAT TAGGGTCGGT     750
AATGGTGGCG ACGACCGTGT AGCTACCGGG CCCCACTGGC GCATGGGTGG     800
AACCGTTGTA GGTAAAGGAG ACGTCAAGCC CCACGGGATG GGTCTCGGCA     850
AGAGCGGCCT TGGGGGTGCC GTCGAAAACG TGTTCCAAAT TGGAGAGCGT     900
GATGGTGGCG GGTGCCTTGA GCACAGTCAC AGAAACAGTG GATTGCACGG     950
GATCGTGCGC TGCCGTGTCT GCAGGTGTGA AGACCACGCT GTAAAAACGG    1000
GTTCCGGCGG ACGGTGCAAG GCCGGACAGG ACAAAGGCAA AGTCGCCGGG    1050
GACGGCGGCT ACTCCGCCGC TCAGGCCGGC CTCCGCAAGG GTTTGCCCGA    1100
AGGTGATGGG TGCGGCTGTG GGCCACATCT CCACAAGGCC GGTGTCCCCC    1150
TCGTCACGCA CCGGCATGAG GGCGGAGAGG AGATGAATGT AACTGGCTTG    1200
GTAATTGATG TCGGGCTCGG TGATTTCCCA TGAGTTCTCC GGCCAAAAAC    1250
CATTCCAATC AAGGTAGGCT TTTTGCACGG GTTGGTCTCG GATCGCCTGA    1300
ATGCTTCCGC TGTATTTGGG CATTGGGACC CGCCCGAAAG AAAACCAGGA    1350
GCGGGACCGT AGAGTGAAGT GAGGGCATTG TCCCAGTCCG GCCATCGCGG    1400
AACCAATGGT GGTAGATTTC ATTGGCTGCA CGGTCAGCGC CGCTGGCATA    1450
CATGTTGCTA AGATAGACCA TGCCCATTGG GTTCACTCCG TGGAGATAGT    1500
GCAGGTAGCC CATCGCGGCA TCGCGATGCG CGGCCGCGTC GGCGGGGTTG    1550
AGCCCAAGCC TCCGTACCCC CTCGAAGAAA AAGCCAGCCT GAGACTTTGT    1600
TTTGTTCGAG CCCCACGTGT AATCCTGATC CTTCAGGTAG GCGCGGTAGG    1650
CGTCGGTCTG GTTATTCCAT GCACCGAGAA ACTCCCCACC GTTTATAGAA    1700
GCCGCCATCC GGTTGCGGAT GTCGGCAGAG ACGCTAGGCG TCGCTCCCGG    1750
GAGGGTCGTG TAGTGGGCGA GAGCTTTTTG TAGCTCACCT TGAAAGGGGA    1800
AGAAATACCA CCACTGCACG GGCTCCATAT CGAGATAGCG CACATCGAAG    1850
AAATCGCGAT AGACCGCACC GCCCGTGCGC TCGAAGAGCA TGGCGGCGGC    1900
CATCACACGG TTGGCTAGCG TATCGTGGGC ATTGCGCGAG GGGCTCACGG    1950
AAGCAAATCC GGTGTTGTCG AAAGGCACAT GAGGATGGAC CATGGTCCAA    2000
TTCCATGCGG CGATGGCAGC GGATTCGAGG GTGACGGCAT AATCGCTCAT    2050
GCCTACGCTC TCAAAGACAG TCGCCCCGAG GGCGAAAGCG GCGGCAGCCA    2100
TGGCAGTGGC CTCGGTCGAG ACGGGGCCGT AGTAACGCGG ATGGGTGTCG    2150
GTGCTCGGCG GGCTGGCGCT CTGGTGCCCC GTCACGGAAA CTTTCCCGAG    2200
AATAGCCCCG CTCGGCTCCT GCATGCGTAA GAGCCAGTCC ATTCCCCATT    2250
TGACTTCGTC AAGCAGGTCG GGGACACCGT TGCCGGATTC CGGGATGCCA    2300
AAATCATCGG TAAAGACGTC AGGCCGCCCT TGATAGGCAA GGAGCAGCTC    2350
CAGGATGACG CGCCCCGTCC ACTCGCTGTA CTTGTTGAAA TCGCCCGCAT    2400
```

Figure 1B

```
CGAACCAACC GCCGCTGAGA TCGCGCTCCA AGGAGGCATT CCCCATATCC        2450
CAGATGGGGC GGCTGGCGAC GTCCTGCGGG TGAGAAGCGG CATCGGCCCA        2500
GTTCGCGTGG GCGTAGGGCA CCTCCTTGGC AAACCCGGAG CGCTGATAGA        2550
AGAACATGCG CACGGCCTCG CGCAGGACAA CATCGTAAAC ATCCGCGCCA        2600
ATGGCGAAAC TATCGGAATG AGTGTTGTTG GCAGGATCGT GGATGCGGTA        2650
GTGGCCGGGC TCGGCAACTA CCGTAAAATC AAACCACCAC ACGCGGTCTC        2700
CCGATTGAAT ATGGATGGCG CCGCCGTTCC ACGGGACCGG TGAGCCGGAG        2750
AAAACCACGA CGCCATCGTT CACGCGACGG ACCTCCAGCG TTGCGCCGGG        2800
GCTGTAGCTC TCGGCGCTGT TCCAGCCAAT CTGCGGGTCG GCGATCACCG        2850
CCACCTTGGT GGCATCGGCG GGGTAACCGA ATTGGTCGAT GCGGATTTTA        2900
TCGGTGTGGG TGGAGGCGAC GAGGGCGGAG CTGCCCATGA GCAGCAAGAA        2950
AAAGCCCGCT GTCGGCCCGA TACCAAAAAA ACGAATAGGG AGAGAAAAAT        3000
TCATAGCAGG ATGTGGATAC GGAAAGGGGG AAAACGGTGC AAAGACCCAA        3050
GCCCAACGCT TGGCGAAAAC TGGATGGTTG GTTTATCAAG AAAAGCGCTT        3100
TTGAGCCAAA AGCTGCGGGC AATCCTTATT GCGTTTCACA ATATTTTCAC        3150
ATCGTCGGCG GCACGACTTT TCGATGGGCG ACTTGACAGC GTATTCTCTC        3200
AGGCGCGAGG CTGCAAACCT TATGAAAAAA GGCCCGCGCA GCGATCTGTC        3250
CCCGGTCAAA ATCCAGTCAA GGTTTGTTCA AGGGTTTGAG GTCTGATAGA        3300
GGCACAGTCG AGCCATCAGC AGTCGCATTG AGTAGGGTTG TTGGAGAAAG        3350
TGTGCAAATG ACCGCTGCCG AAGGAACTGT GGAGACAAAA AGCATATTTT        3400
CCTCGCCAAG                                                   3410
```

Figure 2
The nucleotide sequence of 029cel ORF

```
ATGAATTTTT CTCTCCCTAT TCGTTTTTTT GGTATCGGGC CGACAGCGGG          50
CTTTTTCTTG CTGCTCATGG GCAGCTCCGC CCTCGTCGCC TCCACCCACA         100
CCGATAAAAT CCGCATCGAC CAATTCGGTT ACCCCGCCGA TGCCACCAAG         150
GTGGCGGTGA TCGCCGACCC GCAGATTGGC TGGAACAGCG CCGAGAGCTA         200
CAGCCCCGGC GCAACGCTGG AGGTCCGTCG CGTGAACGAT GGCGTCGTGG         250
TTTTCTCCGG CTCACCGGTC CCGTGGAACG GCGGCGCCAT CCATATTCAA         300
TCGGGAGACC GCGTGTGGTG GTTTGATTTT ACGGTAGTTG CCGAGCCCGG         350
CCACTACCGC ATCCACGATC CTGCCAACAA CACTCATTCC GATAGTTTCG         400
CCATTGGCGC GGATGTTTAC GATGTTGTCC TGCGCGAGGC CGTGCGCATG         450
TTCTTCTATC AGCGCTCCGG GTTTGCCAAG GAGGTGCCCT ACGCCCACGC         500
GAACTGGGCC GATGCCGCTT CTCACCCGCA GGACGTCGCC AGCCGCCCCA         550
TCTGGGATAT GGGGAATGCC TCCTTGGAGC GCGATCTCAG CGGCGGTTGG         600
TTCGATGCGG GCGATTTCAA CAAGTACAGC GAGTGGACGG GGCGCGTCAT         650
CCTGGAGCTG CTCCTTGCCT ATCAAGGGCG GCCTGACGTC TTTACCGATG         700
ATTTTGGCAT CCCGGAATCC GGCAACGGTG TCCCCGACCT GCTTGACGAA         750
GTCAAATGGG GAATGGACTG GCTCTTACGC ATGCAGGAGC CGAGCGGGGC         800
TATTCTCGGG AAAGTTTCCG TGACGGGGCA CCAGAGCGCC AGCCCGCCGA         850
GCACCGACAC CCATCCGCGT TACTACGGCC CCGTCTCGAC CGAGGCCACT         900
GCCATGGCTG CCGCCGCTTT CGCCCTCGGG GCGACTGTCT TTGAGAGCGT         950
AGGCATGAGC GATTATGCCG TCACCCTCGA ATCCGCTGCC ATCGCCGCAT        1000
GGAATTGGAC CATGGTCCAT CCTCATGTGC CTTTCGACAA CACCGGATTT        1050
GCTTCCGTGA GCCCCTCGCG CAATGCCCAC GATACGCTAG CCAACCGTGT        1100
GATGGCCGCC GCCATGCTCT TCGAGCGCAC GGGCGGTGCG GTCTATCGCG        1150
ATTTCTTCGA TGTGCGCTAT CTCGATATGG AGCCCGTGCA GTGGTGGTAT        1200
TTCTTCCCCT TTCAAGGTGA GCTACAAAAA GCTCTCGCCC ACTACACGAC        1250
CCTCCCGGGA GCGACGCCTA GCGTCTCTGC CGACATCCGC AACCGGATGG        1300
CGGCTTCTAT AAACGGTGGG GAGTTTCTCG GTGCATGGAA TAACCAGACC        1350
GACGCCTACC GCGCCTACCT GAAGGATCAG GATTACACGT GGGGCTCGAA        1400
CAAAACAAAG TCTCAGGCTG GCTTTTTCTT CGAGGGGGTA CGGAGGCTTG        1450
GGCTCAACCC CGCCGACGCG GCCGCGCATC GCGATGCCGC GATGGGCTAC        1500
CTGCACTATC TCCACGGAGT GAACCCAATG GGCATGGTCT ATCTTAGCAA        1550
CATGTATGCC AGCGGCGCTG ACCGTGCAGC CAATGAAATC TACCACCATT        1600
GGTTCCGCGA TGGCCGGACT GGGACAATGC CCTCACTTCA CTCTACGGTC        1650
CCGCTCCTGG TTTTCTTTCG GGCGGGTCCC AATGCCCAAA TACAGCGGAA        1700
GCATTCAGGC GATCCGAGAC CAACCCGTGC AAAAAGCCTA CCTTGA            1746
```

Figure 3
The translated 029cel protein composed of 581 amino acids

```
MNFSLPIRFF GIGPTAGFFL LLMGSSALVA STHTDKIRID QFGYPADATK      50
VAVIADPQIG WNSAESYSPG ATLEVRRVND GVVVFSGSPV PWNGGAIHIQ     100
SGDRVWWFDF TVVAEPGHYR IHDPANNTHS DSFAIGADVY DVVLREAVRM     150
FFYQRSGFAK EVPYAHANWA DAASHPQDVA SRPIWDMGNA SLERDLSGGW     200
FDAGDFNKYS EWTGRVILEL LLAYQGRPDV FTDDFGIPES GNGVPDLLDE     250
VKWGMDWLLR MQEPSGAILG KVSVTGHQSA SPPSTDTHPR YYGPVSTEAT     300
AMAAAAFALG ATVFESVGMS DYAVTLESAA IAAWNWTMVH PHVPFDNTGF     350
ASVSPSRNAH DTLANRVMAA AMLFERTGGA VYRDFFDVRY LDMEPVQWWY     400
FFPFQGELQK ALAHYTTLPG ATPSVSADIR NRMAASINGG EFLGAWNNQT     450
DAYRAYLKDQ DYTWGSNKTK SQAGFFFEGV RRLGLNPADA AAHRDAAMGY     500
LHYLHGVNPM GMVYLSNMYA SGADRAANEI YHHWFRDGRT GTMPSLHSTV     550
PLLVFFRAGP NAQIQRKHSG DPRPTRAKSL P                         581
```